Protein Design Platform: Protease Module

Find appropriate backbone geometry

Substitute side chains

Validate substitutions

Input into database
Input into program

Public Database

engeneOS Database

Computational procedure/program

db1 Reference structures

p1 Parameters of active site?

db2 Active Site Geometry

db6 SwissProt

db7 PDB

db8 Enzyme

db9 Rotamers

db3 Proteins to use as framework (Protein ID)

Select Targets

p2 Where are the Triad Pockets?

db4 Potential Triad Sites by Protein ID

p3 Mutate pockets into triads

db5 Mutated Proteins (.pdb)

p4 Engineered Side Chain Compatible?

p5 Engineered Triad form H-bond?

p6 Ligand Accessible?

p7 Engineered Triad structurally compatible to natural triads?

annotation

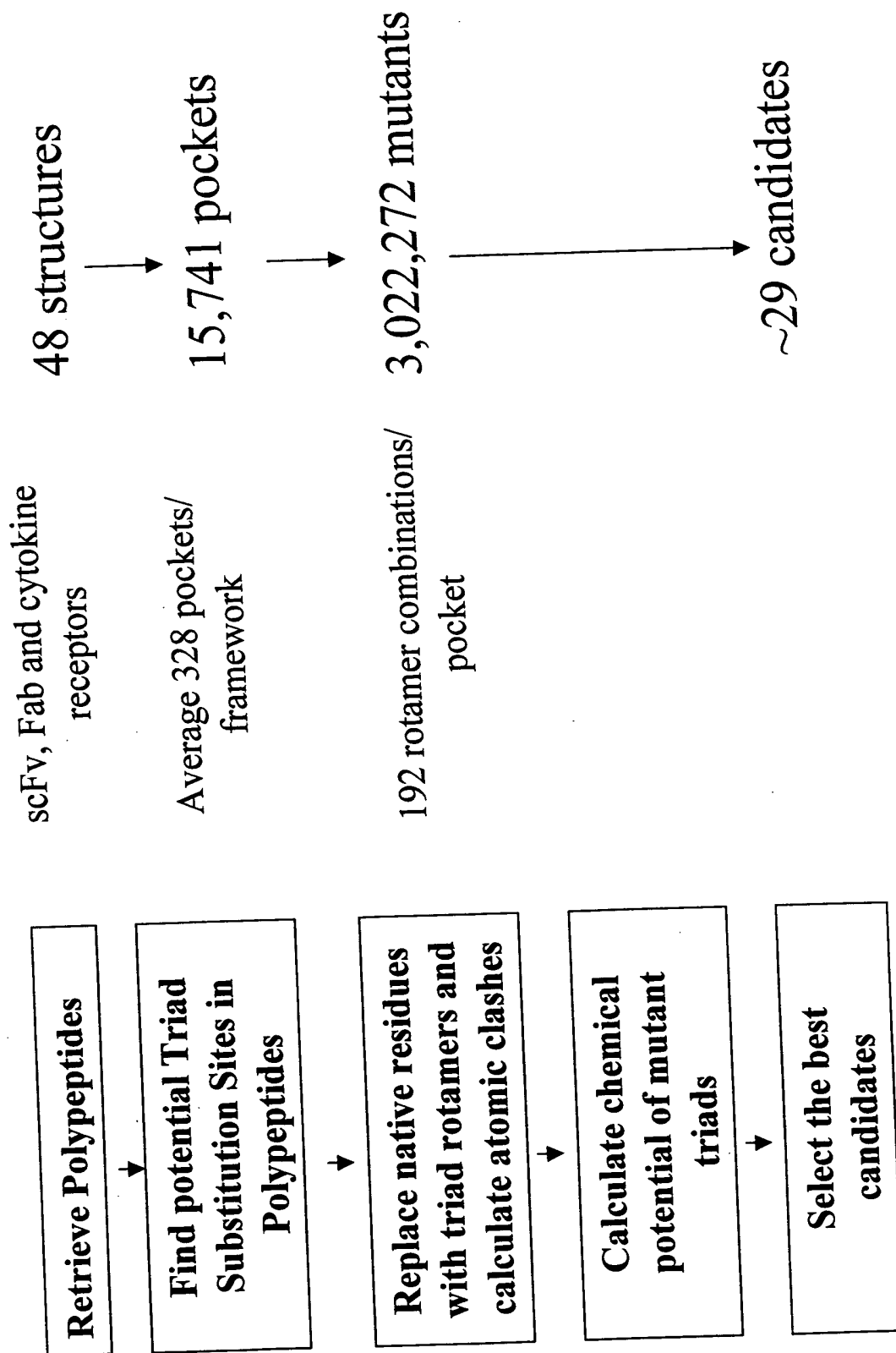db10 Protein Annotation

db11 References

IF1 Workstation

Select Mutants

Figure 1

1/3

Triad Design Schematic

48 structures

↓

15,741 pockets

↓

3,022,272 mutants

→

~29 candidates

scFv, Fab and cytokine receptors

Average 328 pockets/ framework

192 rotamer combinations/ pocket

Retrieve Polypeptides

↓

Find potential Triad Substitution Sites in Polypeptides

↓

Replace native residues with triad rotamers and calculate atomic clashes

↓

Calculate chemical potential of mutant triads

↓

Select the best candidates

Figure 2

FIGURE 3

Overall mutation: 1TNF_b_tyr_119_asp, c_tyr_119_his, c_tyr_59_ser

5

1. Chain B mutation: b_tyr_119_asp

>1TNF_alpha
GTCAGATCATCTTCTCGAACCCCGAGTGACAAGCCTGTAGCCCATGTTGTAGCAAACCCTCAAGCTGAG
GGGCAGCTCCAGTGGCTGAACCGCCGGGCCAATGCCCTCCTGGCCAATGGCGTGGAGCTGAGAGATAAC
10 CAGCTGGTGGTGCCATCAGAGGGCCTGTACCTCATCTACTCCCAGGTCCTCTTCAAGGGCCAAGGCTGC
CCCTCCACCCATGTGCTCCTCACCCACACCATCAGCCGCATCGCCGTCTCCTACCAGACCAAGGTCAAC
CTCCTCTCTGCCATCAAGAGCCCCTGCCAGAGGGAGACCCCAGAGGGGGCTGAGGCCAAGCCCTGGTAT
GAGCCCATCGATCTGGGAGGGGTCTTCCAGCTGGAGAAGGGTGACCGACTCAGCGCTGAGATCAATCGG
CCCGACTATCTCGACTTTGCCGAGTCTGGGCAGGTCTACTTTGGGATCATTGCCCTGTGA

15 Resulting Protein sequence, chain B:
VRSSSRTPSDKPVAHVVANPQAEGQLQWLNRRANALLANGVELRDNQLVVPSEGLYLIYS
QVLFKGQGCPSTHVLLTHTISRIAVSYQTKVNLLSAIKSPCQRETPEGAEAKPWYEPIDL
GGVFQLEKGDRLSAEINRPDYLLFAESGQVYFGIIAL

20 2. Chain C mutations: c_tyr_119_his, c_tyr_59_ser

>1TNF_alpha
GTCAGATCATCTTCTCGAACCCCGAGTGACAAGCCTGTAGCCCATGTTGTAGCAAACCCTCAAGCTGAG
GGGCAGCTCCAGTGGCTGAACCGCCGGGCCAATGCCCTCCTGGCCAATGGCGTGGAGCTGAGAGATAAC
CAGCTGGTGGTGCCATCAGAGGGCCTGTACCTCATCAGTTCCCAGGTCCTCTTCAAGGGCCAAGGCTGC
25 CCCTCCACCCATGTGCTCCTCACCCACACCATCAGCCGCATCGCCGTCTCCTACCAGACCAAGGTCAAC
CTCCTCTCTGCCATCAAGAGCCCCTGCCAGAGGGAGACCCCAGAGGGGGCTGAGGCCAAGCCCTGGTAT
GAGCCCATCCATCTGGGAGGGGTCTTCCAGCTGGAGAAGGGTGACCGACTCAGCGCTGAGATCAATCGG
CCCGACTATCTCGACTTTGCCGAGTCTGGGCAGGTCTACTTTGGGATCATTGCCCTGTGA

Resulting Protein sequence, chain C:
30 VRSSSRTPSDKPVAHVVANPQAEGQLQWLNRRANALLANGVELRDNQLVVPSEGLYLISS
QVLFKGQGCPSTHVLLTHTISRIAVSYQTKVNLLSAIKSPCQRETPEGAEAKPWYEPIHL
GGVFQLEKGDRLSAEINRPDYLLFAESGQVYFGIIAL